

Sep. 22, 1998

Digital Library Technologies and Information Infrastructure

Soon J. Hyun

School of Information and Computer Engineering
Information and Communications University

ICU

shyun@icu.ac.kr, <http://www.icu.ac.kr>

ICU: Information and Communications Univ.

Digital Library

Definitions;

A library that has been extended and enhanced by the application of digital technology.

Contents and services accessible and deliverable via electronic networks to and from the worldwide libraries and commercial information resources.

A large, organized collection of multimedia information for providing users of a community with coherent access to certain information services.

A realistic means to build an Information Infrastructure for advanced Internet services of the future.

more ...

Historically,

Computer-based Library, Library of the Future, Library Automation (1965, Licklider)

Hypertext and Hyperspace (1974, nelson)

Paperless Library (1978, Lancaster)

Electronic Library, Virtual Library, Library without Wall (1980s-90s)

Digital Library (DLI, 1994)

3

DL & KM '98

Data, Information and Knowledge

- Known facts or values that can be physically recorded in the form of data objects
- Meaning of those facts or values as understood by some users
- Value-added information about data objects useful for decision-support in design, planning, manufacturing, and business functions

4

DL & KM '98

Ever-increasing Needs for Digital Information

- **Web users explosion**

- > 17 M web servers (97)
- > Number of users doubles every 6 months
- > Network bandwidth triples every year

- **Profusion of digital tools**

- > Software: multimedia applications, CAD, satellite image, medical images, ...
- > Hardware: digital camera and videocam, scanners, DVD,...

- **Growing demand for increased amount of cable content**

- > moving to digital TV (by 2005)

- **Business communities need to harness new presentation tech**

- > video conference, virtual enterprise, virtual office, electronic commerce, ...

- **Decrease in digital storage cost**

- > magnetic, optical and tape
- > cost of digital storage decreases 40% /year

5

DL & KM '98

A New Scale for Storage

1 TERABYTE = 1,000 Gigabytes = 1,000,000 Megabyte
= 300,000,000 pages of text
= 20,000,000 pages of scanned documents
= 1,000,000 pages of scanned color images
= 100,000 medical X-rays
= 1,800 hours of CD quality audio
= 500 hours of good quality video
= 250 movies

PETA = 1,000 TERA = 1,000,000 GIGA

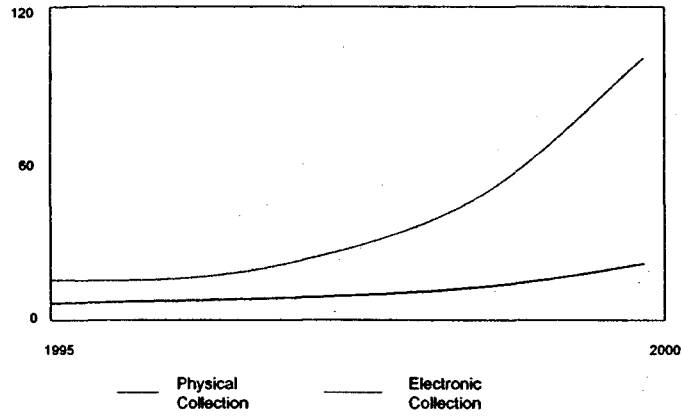
EXE = 1,000 PETA = 1,000,000 TERA

6

DL & KM '98

Digital Collection

Holdings (in Mil.)



7

DL & KM '98

Information services require

- Computing Power
- Network Bandwidth
- Information Management Skill

Information services generally do

- Digitizing
- Delivery
- Display

8

DL & KM '98

Web vs. Digital Library

- Information and Knowledge are not organized nor properly protected.
- Too much burden for users in finding information (still, not guaranteed to be reliable).
- No content search of multimedia information.
- Low applicability to other advanced services
- Slow delivery of multimedia information

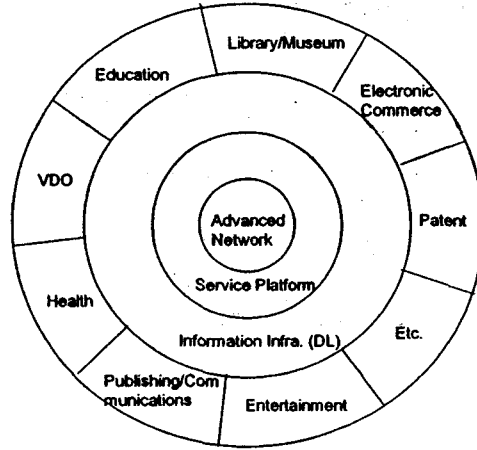
Information Infrastructure

A common application-sharing construct which enables both information consumers and information providers of a community to easily use the information services.

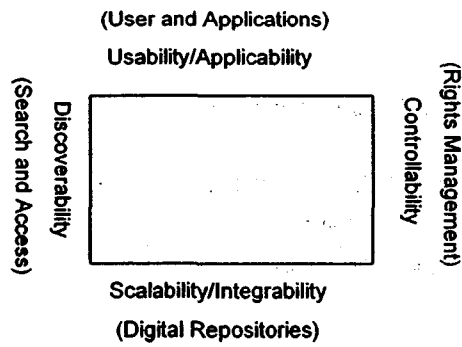
Communication Infra and Information Infra

Application 1	Application 2	Application 3	...
Information Infra			
Communication Infra			

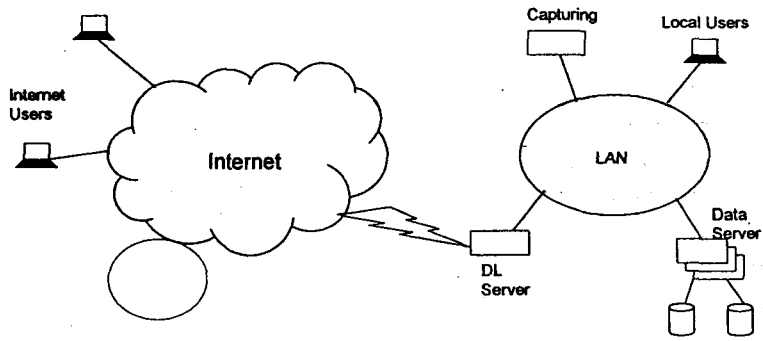
Information Infrastructure: Virtual Information Network



Conceptual Design



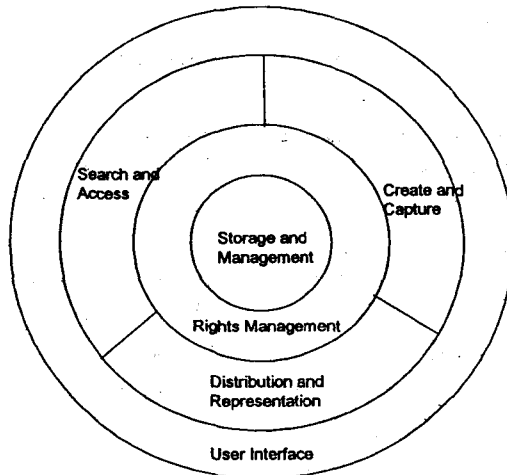
General Configuration



13
DL & KM '98

Technologies

Research and Development Construct



14
DL & KM '98

Research and Development Issues

- Collection and Organization
- Resource Discovery
- User Interface
- Legacy Integration
- Interoperability
- Multimedia and Networking
- Rights Management
- Object-oriented Data Management and Modeling
- Standards
- Advanced Service Development
- Social, Economic, and Legal Issues

15

DL & KM '98

Resource Discovery and User Interface

Content-based Search

- Full text search
 - Given some keywords a full-text search engine returns useful statistics.
- Image content search
 - How do you find similar pictures in a database of million pictures?
 - Image query with shape, average color, color mixture, texture, color position, etc.
- Multi-search
 - Combined parametric, full text and image search:
e.g., Color/pattern + "Made in Korea" (Full-text) + "type=suit, price<\$100" (Catalog search)
- Video content search
 - Automatically detect scene changes: % change from frame to frame
 - Uses Image search technology: locates frame to meet search criteria
- Speech recognition
 - Software for visually handicapped

16

DL & KM '98

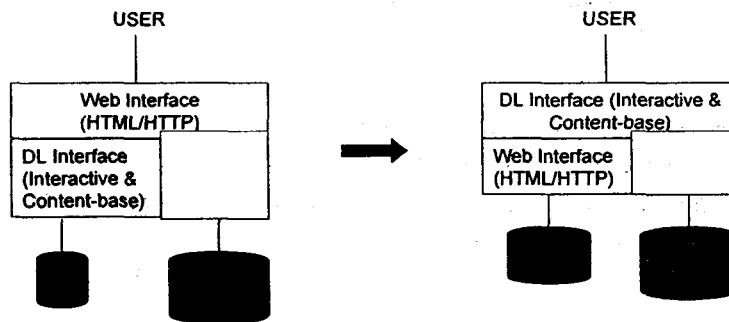
User Interface

- User-centered Service vs. System-centered Service
 - >> Interactive Querying
 - >> Content-based Querying
 - >> Active Querying Environment
 - >> 3D Interface
- Information Outlining
 - >> Feature Extraction and Analysis: Frequencies, Singularity, Topics, ...
 - >> Visualization in Multiple Views: Graphs, Diagrams, Geographical views, Topic extraction ...
 - >> Interaction and Navigation with Object picking, Sub/super-setting, ...
- Web Translation
 - >> English Web Pages >> Personal Proxy
 - >> Translation Engine (HTML Layout Analysis>>Source language analysis>>Source-to-target transfer >> Target language generation >> HTML Layout generation)
 - >> Personal proxy >> Web browser

17

DL & KM '98

User Interface: migration



18

DL & KM '98

Legacy Integration, Interoperability

- Heterogeneous Information Systems
- Interoperation independent of sources
- data model (schema), language, attribute naming, data structure, ...
- Federation, moderation
- Object-oriented brokering

19

DL & KM '98

Multimedia

multi-modes of communication

Text, sound, graphics, image, video, animation, music, ...

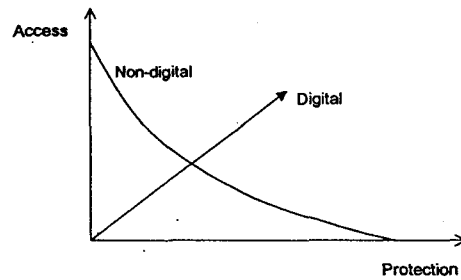
- Text
- Audio: waveform, MIDI, sampling, compression, filtering,...
- Animation: Sprites, tweening, color cycling; Morphing, cell animation, 3D, ...
- Image: video capturing, scanning, digital camera, photo CD, compression(JPEG), ...
- Video: Standards, compression, MPEG-1, MPEG-2, ...

20

DL & KM '98

Rights Management

- Copyright Protection
- Marking: Visible/invisible watermarking
- Authentication
- Encryption
- Charging and Payment: metering and billing services, royalty management



21

DL & KM '98

Standards

- document representation
Markup languages
- metadata protocol
Z39.50

Social, Economic, and Legal Issues

- Intellectual Property
- Copyright
- Taxation
- etc.

22

DL & KM '98

Application/Service Development

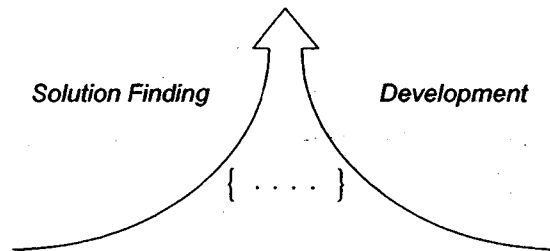
- Government Information Locator Service
- Electronic Commerce
- Electronic Business
- Virtual Enterprise
- Virtual Digital Office
- Electronic Publishing
- Health
- Patent
- Digital CATV, VOD
- Geographic Information System
- Environmental Service
- Remote Education, Cyber University
- User-make applications, ...

23

DL & KM '98

Developing DL Technologies

{ Applications | Dist. Learning, GILS, GIS, Library/Museum, CALS, Patent, EC, etc. }



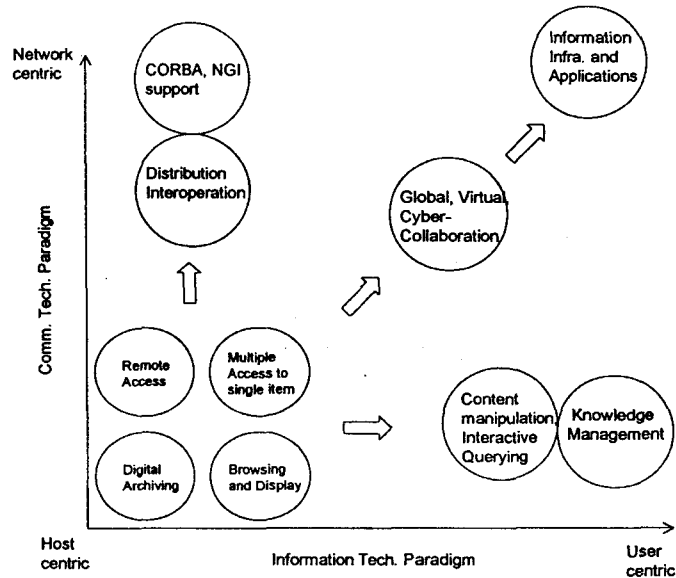
{ DLTechnologies | DBMS, Multimedia, Internet, Search Mgt, Rights Management, Cataloging, Capturing, Modeling, Z39.50, Migration, Setup, etc. }

24

DL & KM '98

Technology Trend

ICU: Information and Communications Univ.



Digital Library: Activities and Operations

ICU: Information and Communications Univ.

International, Asia-pacific

- Workshop on Digital Library, May 1994
- DL/NSF-ARPA-NASA, Sep. 1994 ~98
- DL12, Fall 1998 ~2002
- ACM special issues on DL, April 1994
- ACM Conference on Digital Libraries
- D-lib Magazine: www.dlib.org
- Vatican Library, NAIST, ...

Domestic

- Int'l Conf. on Digital Libraries and Info. Services, Seoul, Sep. 1996
- Digital Library and Knowledge Management Conference, Seoul, Sep. 1998
- Digital Library Workshop, DPC, Seoul, scheduled in November, 1998
- Establishment of APAN's DL working group
- 5 Consortium Libraries, ETLARS, KORDIC, MIRAGE, ...

25
DL & KM '98

APAN-KR's Digital Library Working Group

- A working group for promoting the research and development of DL-based information infrastructure for APAN, Korea.
- Initiating Members
 - Kwangwoon Univ., SNU, Chungnam Univ., Yonsei Univ., Sookmyung Univ., Catholic Univ., and ICU
 - More members are expected to join
 - For more information: <http://www.apan.net>, shyun@icu.ac.kr
- To officially form a DLWG/APAN among APAN member countries in the next meeting
- To establish a linkage with other regional consortiums